Variational Methods for Latent Variable Problems

Ryan Giordano (for Johns Hopkins Biostats BLAST working group)

Oct, 2021

Massachusetts Institute of Technology

Outline

Outline for today:

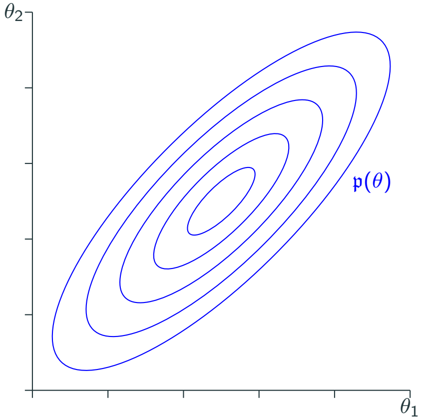
- Some examples of latent variable models
- A template: The Neyman-Scott “paradox” and marginalization
- Bayesian versus frequentist approaches to marginalization
- The classical EM algorithm (in brief)

Next week, we will build on these ideas to present more general variational inference.

KL divergence exercises

$$\mathcal{Q} = \{\text{Bivariate normals}\}$$

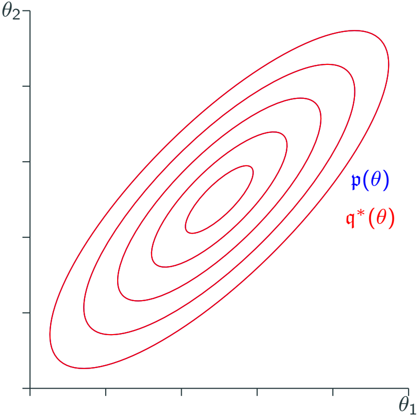
$$\mathfrak{q}^*(\theta) = \underset{\mathfrak{q}\in\mathcal{Q}}{\operatorname{argmin}} \operatorname{KL}\left(\mathfrak{q}(\theta) || \mathfrak{p}(\theta)\right)$$



KL divergence exercises

$$\mathcal{Q} = \{\text{Bivariate normals}\}$$

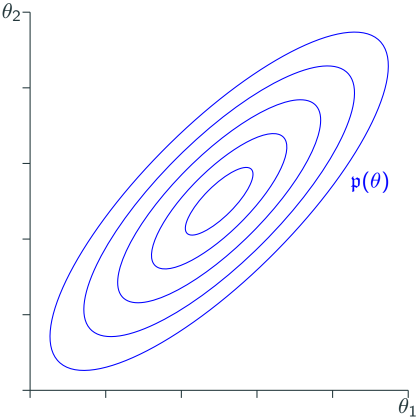
$$\mathfrak{q}^*(\theta) = \underset{\mathfrak{q} \in \mathcal{Q}}{\operatorname{argmin}} \operatorname{KL}\left(\mathfrak{q}(\theta) || \mathfrak{p}(\theta)\right)$$



KL divergence exercises

$$\mathcal{Q} = \{\text{Independent bivariate normals}\}$$

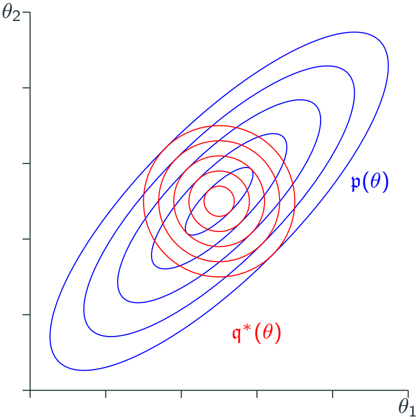
$$\mathfrak{q}^*(\theta) = \underset{\mathfrak{q}\in\mathcal{Q}}{\operatorname{argmin}} \, \mathrm{KL}\left(\mathfrak{q}(\theta)||\mathfrak{p}(\theta)\right)$$



KL divergence exercises

$$\mathcal{Q} = \{\text{Independent bivariate normals}\}$$

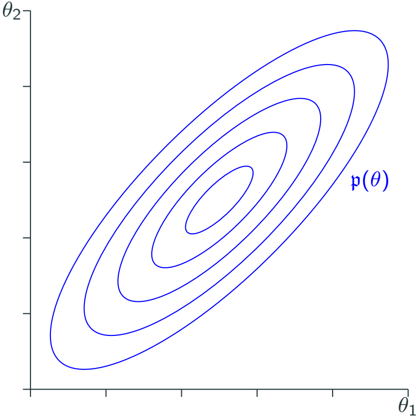
$$\mathfrak{q}^*(\theta) = \operatorname*{argmin}_{\mathfrak{q}\in\mathcal{Q}} \mathrm{KL}\left(\mathfrak{q}(\theta)||\mathfrak{p}(\theta)\right)$$



KL divergence exercises

$$\mathcal{Q} = \{\text{Independent bivariate normals}\}$$

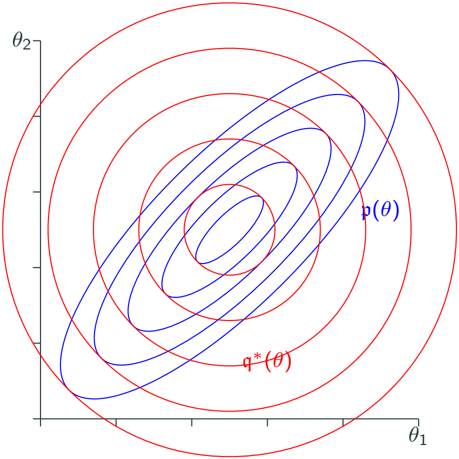
$$\mathfrak{q}^*(\theta) = \underset{\mathfrak{q}\in\mathcal{Q}}{\operatorname{argmin}} \mathrm{KL}\left(\mathfrak{p}(\theta)||\mathfrak{q}(\theta)\right)$$



KL divergence exercises

$$\mathcal{Q} = \{\text{Independent bivariate normals}\}$$

$$\mathfrak{q}^*(\theta) = \operatorname*{argmin}_{\mathfrak{q} \in \mathcal{Q}} \mathrm{KL}\left(\mathfrak{p}(\theta) || \mathfrak{q}(\theta)\right)$$



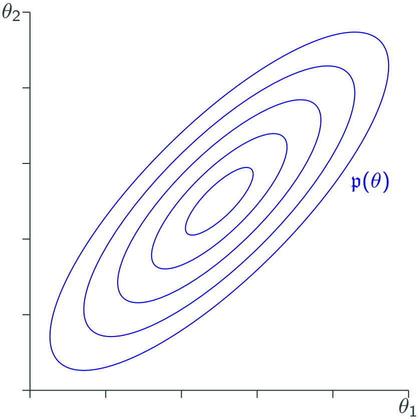
KL divergence exercises

Recall that

$$\mathrm{KL}\left(\mathfrak{q}(\theta)||\mathfrak{p}(\theta)\right) = -\underset{\mathfrak{q}(\theta)}{\mathbb{E}}\left[\log \mathfrak{p}(\theta)\right] + \underset{\mathfrak{q}(\theta)}{\mathbb{E}}\left[\log \mathfrak{q}(\theta)\right]$$

$$\mathcal{Q} = \{\text{Bivariate normals}\}$$

$$\mathfrak{q}^*(\theta) = \underset{\mathfrak{q}\in\mathcal{Q}}{\operatorname{argmin}}\left(-\underset{\mathfrak{q}(\theta)}{\mathbb{E}}\left[\log \mathfrak{p}(\theta)\right] + \cancel{\underset{\mathfrak{q}(\theta)}{\mathbb{E}}\left[\log \mathfrak{q}(\theta)\right]}\right)$$



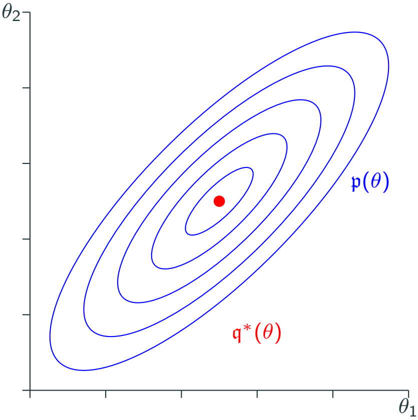
KL divergence exercises

Recall that

$$\mathrm{KL}\left(\mathfrak{q}(\theta)||\mathfrak{p}(\theta)\right) = -\underset{\mathfrak{q}(\theta)}{\mathbb{E}}[\log \mathfrak{p}(\theta)] + \underset{\mathfrak{q}(\theta)}{\mathbb{E}}[\log \mathfrak{q}(\theta)]$$

$$\mathcal{Q} = \{\text{Bivariate normals}\}$$

$$\mathfrak{q}^*(\theta) = \underset{\mathfrak{q}\in\mathcal{Q}}{\operatorname{argmin}}\left(-\underset{\mathfrak{q}(\theta)}{\mathbb{E}}[\log \mathfrak{p}(\theta)] + \cancel{\underset{\mathfrak{q}(\theta)}{\mathbb{E}}[\log \mathfrak{q}(\theta)]}\right)$$



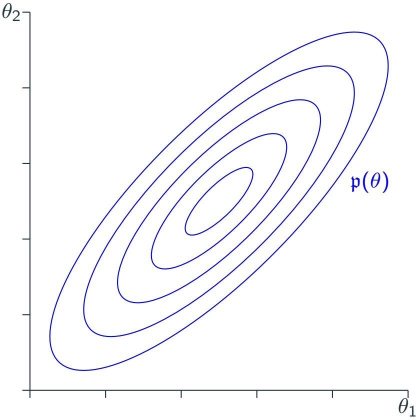
KL divergence exercises

Recall that

$$\mathrm{KL}\left(\mathfrak{q}(\theta)||\mathfrak{p}(\theta)\right) = - \underset{\mathfrak{q}(\theta)}{\mathbb{E}}\left[\log \mathfrak{p}(\theta)\right] + \underset{\mathfrak{q}(\theta)}{\mathbb{E}}\left[\log \mathfrak{q}(\theta)\right]$$

$$\mathcal{Q} = \{\text{Bivariate normals}\}$$

$$\mathfrak{q}^*(\theta) = \underset{\mathfrak{q}\in\mathcal{Q}}{\operatorname{argmin}}\left(\cancel{- \underset{\mathfrak{q}(\theta)}{\mathbb{E}}\left[\log \mathfrak{p}(\theta)\right]} + \underset{\mathfrak{q}(\theta)}{\mathbb{E}}\left[\log \mathfrak{q}(\theta)\right]\right)$$



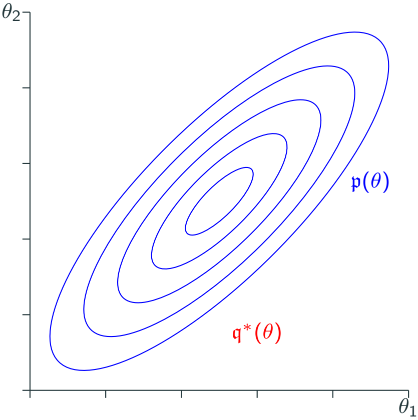
KL divergence exercises

Recall that

$$\mathrm{KL}\left(\mathfrak{q}(\theta)||\mathfrak{p}(\theta)\right) = - \mathop{\mathbb{E}}_{\mathfrak{q}(\theta)}\left[\log \mathfrak{p}(\theta)\right] + \mathop{\mathbb{E}}_{\mathfrak{q}(\theta)}\left[\log \mathfrak{q}(\theta)\right]$$

$$\mathcal{Q} = \{\text{Bivariate normals}\}$$

$$\mathfrak{q}^*(\theta) = \operatorname*{argmin}_{\mathfrak{q}\in\mathcal{Q}}\left(- \cancel{\mathop{\mathbb{E}}_{\mathfrak{q}(\theta)}\left[\log \mathfrak{p}(\theta)\right]} + \mathop{\mathbb{E}}_{\mathfrak{q}(\theta)}\left[\log \mathfrak{q}(\theta)\right]\right)$$



Conclusions